

Re Run

#11

ENTERED



OIPE

## RAW SEQUENCE LISTING

DATE: 05/09/2003

PATENT APPLICATION: US/09/870,353A

TIME: 14:31:42

Input Set : N:\Crf4\05062003\I870353A.raw

Output Set: N:\CRF4\05072003\I870353A.raw

1 <110> APPLICANT: Wang, Yan  
 2 Xi, Lei  
 3 Prosen, Dennis E.  
 4 MJ Bioworks, Inc.  
 5 <120> TITLE OF INVENTION: Improved Nucleic Acid Modifying Enzymes  
 6 <130> FILE REFERENCE: 020130-000111US  
 C--> 7 <140> CURRENT APPLICATION NUMBER: US/09/870,353A  
 8 <141> CURRENT FILING DATE: 2001-05-30  
 9 <150> PRIOR APPLICATION NUMBER: US 60/207,567  
 10 <151> PRIOR FILING DATE: 2000-05-26  
 11 <150> PRIOR APPLICATION NUMBER: US 09/640,958  
 12 <151> PRIOR FILING DATE: 2000-08-16  
 13 <160> NUMBER OF SEQ ID NOS: 34  
 14 <170> SOFTWARE: PatentIn Ver. 2.1  
 16 <210> SEQ ID NO: 1  
 17 <211> LENGTH: 189  
 18 <212> TYPE: DNA  
 19 <213> ORGANISM: Artificial Sequence  
 20 <220> FEATURE:  
 21 <223> OTHER INFORMATION: Description of Artificial Sequence:synthetic  
 22 sequence non-specific double-stranded nucleic acid  
 23 binding protein Ssod7d  
 24 <220> FEATURE:  
 25 <221> NAME/KEY: CDS  
 26 <222> LOCATION: (1)..(189)  
 27 <223> OTHER INFORMATION: Ssod7d  
 28 <400> SEQUENCE: 1  
 29 gcaaccgtaa agttcaagta caaaggcgaa gaaaaagagg tagacatctc caagatcaag 60  
 30 aaagtatggc gtgtgggcaa gatgatctcc ttcacctacg acgagggcgg tggcaagacc 120  
 31 ggccgtggtg cggtaagcga aaaggacgcg ccgaaggagc tgctgcagat gctggagaag 180  
 32 cagaaaaag 189  
 34 <210> SEQ ID NO: 2  
 35 <211> LENGTH: 63  
 36 <212> TYPE: PRT  
 37 <213> ORGANISM: Artificial Sequence  
 38 <220> FEATURE:  
 39 <223> OTHER INFORMATION: Description of Artificial Sequence:synthetic  
 40 sequence non-specific double-stranded nucleic acid  
 41 binding protein Ssod7d  
 42 <400> SEQUENCE: 2  
 43 Ala Thr Val Lys Phe Lys Tyr Lys Gly Glu Glu Lys Glu Val Asp Ile  
 44 1 5 10 15  
 45 Ser Lys Ile Lys Lys Val Trp Arg Val Gly Lys Met Ile Ser Phe Thr

## RAW SEQUENCE LISTING

DATE: 05/09/2003

PATENT APPLICATION: US/09/870,353A

TIME: 14:31:42

Input Set : N:\Cr4\05062003\I870353A.raw

Output Set: N:\CRF4\05072003\I870353A.raw

```

46          20          25          30
47      Tyr Asp Glu Gly Gly Gly Lys Thr Gly Arg Gly Ala Val Ser Glu Lys
48          35          40          45
49      Asp Ala Pro Lys Glu Leu Leu Met Leu Glu Lys Gln Lys Lys
50          50          55          60
52 <210> SEQ ID NO: 3
53 <211> LENGTH: 1899
54 <212> TYPE: DNA
55 <213> ORGANISM: Artificial Sequence
56 <220> FEATURE:
57 <223> OTHER INFORMATION: Description of Artificial Sequence:fusion protein
58      Ssod7d-deltaTaq
59 <220> FEATURE:
60 <221> NAME/KEY: CDS
61 <222> LOCATION: (1)..(1899)
62 <223> OTHER INFORMATION: Ssod7d-deltaTaq
63 <400> SEQUENCE: 3
64      atgattacga attcgagcgc aaccgtaaag ttcaagtaca aaggcgaaga aaaagaggta 60
65      gacatctcca agatcaagaa agtatggcgt gtgggcaaga tgatctcctt cacctacgac 120
66      gagggcgggtg gcaagaccgg ccgtggtgcg gtaagcgaaa aggacgcgcc gaaggagctg 180
67      ctgcagatgc tggagaagca gaaaaagggc ggcggtgtca ctagtcccaa ggccctggag 240
68      gagggccctt ggccccgcgc ggaaggggcc ttctgtgggt ttgtgctttc ccgcaaggag 300
69      cccatgtggg ccgatcttct ggccctggcc gccgccaggg ggggccgggt ccaccgggcc 360
70      cccgagcctt ataaagccct cagggacctg aaggaggcgc gggggcttct cgccaaagac 420
71      ctgagcgttc tggccctgag ggaaggcctt ggccctccgc ccggcgacga ccccatgctc 480
72      ctgcgcctacc tcctggacct ttccaacacc acccccagag ggggtggccc gcgctacggc 540
73      ggggagtgga cggaggaggc gggggagcgg gccgcccttt ccgagaggct cttcgccaac 600
74      ctgtggggga ggcttgaggg ggaggagagg ctectttggc ttaccggga ggtggagagg 660
75      cccctttccg ctgtcctggc ccacatggag gccacggggg tgcgcctgga cgtggcctat 720
76      ctcagggcct tgtccctgga ggtggccgag gagatcgccc gcctcgaggc cgaggtcttc 780
77      cgcctggccg gccacccctt caacctcaac tcccgggacc agctggaaa ggtcctcttt 840
78      gacgagctag ggcttcccgc catcggcaag acggagaaga ccggcaagcg ctccaccagc 900
79      gccgccgtcc tggaggccct ccgcgaggcc caccctatcg tggagaagat cctgcagtac 960
80      cgggagctca ccaagctgaa gagcacctac attgacctt tgccggacct catccacccc 1020
81      aggacgggcc gcctccacac ccgcttcaac cagacggcca cggccacggg caggctaagt 1080
82      agctccgatc ccaacctcca gaacatcccc gtccgcaccc cgcttgggca gaggatccgc 1140
83      cgggccttca tcgccgagga ggggtggcta ttggtggccc tggactatag ccagatagag 1200
84      ctcagggtgc tggcccacct ctccggcgac gagaacctga tccgggtctt ccaggagggg 1260
85      cgggacatcc acacggagac cgccagctgg atgttcggcg tccccggga ggcctggac 1320
86      cccctgatgc gccggggcgc caagaccatc aacttcgggg tcctctacgg catgtcggcc 1380
87      caccgcctct cccaggagct agccatccct tacgaggagg cccaggcctt cattgagcgc 1440
88      tactttcaga gcttccccaa ggtgcggggc tggattgaga agaccctgga ggagggcagg 1500
89      aggcgggggg acgtggagac cctcttcggc cgccgccgct acgtgccaga cctagaggcc 1560
90      cgggtgaaga gcgtgcggga ggcggccgag cgcattggct tcaacatgcc cgtccagggc 1620
91      accgccgccg acctcatgaa gctggctatg gtgaagctct tccccaggct ggaggaaatg 1680
92      ggggccagga tgctccttca ggtccacgac gagctggtcc tcgaggcccc aaaagagagg 1740
93      gcggaggccg tggcccggct ggccaaggag gtcatggagg ggggtgtatcc cctggccgtg 1800
94      cccctggagg tggaggtggg gataggggag gactggctct ccgccaagga gggcattgat 1860
95      ggccgcggcg gaggcgggca tcatcatcat catcattaa 1899

```

## RAW SEQUENCE LISTING

DATE: 05/09/2003

PATENT APPLICATION: US/09/870,353A

TIME: 14:31:42

Input Set : N:\CrF4\05062003\I870353A.raw

Output Set: N:\CRF4\05072003\I870353A.raw

```

97 <210> SEQ ID NO: 4
98 <211> LENGTH: 632
99 <212> TYPE: PRT
100 <213> ORGANISM: Artificial Sequence
101 <220> FEATURE:
102 <223> OTHER INFORMATION: Description of Artificial Sequence:fusion protein
103     Ssod7d-deltaTaq
104 <400> SEQUENCE: 4
105     Met Ile Thr Asn Ser Ser Ala Thr Val Lys Phe Lys Tyr Lys Gly Glu
106         1             5             10             15
107     Glu Lys Glu Val Asp Ile Ser Lys Ile Lys Lys Val Trp Arg Val Gly
108             20             25             30
109     Lys Met Ile Ser Phe Thr Tyr Asp Glu Gly Gly Gly Lys Thr Gly Arg
110             35             40             45
111     Gly Ala Val Ser Glu Lys Asp Ala Pro Lys Glu Leu Leu Gln Met Leu
112             50             55             60
113     Glu Lys Gln Lys Lys Gly Gly Gly Val Thr Ser Pro Lys Ala Leu Glu
114             65             70             75             80
115     Glu Ala Pro Trp Pro Pro Glu Gly Ala Phe Val Gly Phe Val Leu
116             85             90             95
117     Ser Arg Lys Glu Pro Met Trp Ala Asp Leu Leu Ala Leu Ala Ala
118             100            105            110
119     Arg Gly Gly Arg Val His Arg Ala Pro Glu Pro Tyr Lys Ala Leu Arg
120             115            120            125
121     Asp Leu Lys Glu Ala Arg Gly Leu Leu Ala Lys Asp Leu Ser Val Leu
122             130            135            140
123     Ala Leu Arg Glu Gly Leu Gly Leu Pro Pro Gly Asp Asp Pro Met Leu
124     145-----150-----155-----160-----
125     Leu Ala Tyr Leu Leu Asp Pro Ser Asn Thr Thr Pro Glu Gly Val Ala
126             165            170            175
127     Arg Arg Tyr Gly Gly Glu Trp Thr Glu Glu Ala Gly Glu Arg Ala Ala
128             180            185            190
129     Leu Ser Glu Arg Leu Phe Ala Asn Leu Trp Gly Arg Leu Glu Gly Glu
130             195            200            205
131     Glu Arg Leu Leu Trp Leu Tyr Arg Glu Val Glu Arg Pro Leu Ser Ala
132             210            215            220
133     Val Leu Ala His Met Glu Ala Thr Gly Val Arg Leu Asp Val Ala Tyr
134             225            230            235            240
135     Leu Arg Ala Leu Ser Leu Glu Val Ala Glu Glu Ile Ala Arg Leu Glu
136             245            250            255
137     Ala Glu Val Phe Arg Leu Ala Gly His Pro Phe Asn Leu Asn Ser Arg
138             260            265            270
139     Asp Gln Leu Glu Arg Val Leu Phe Asp Glu Leu Gly Leu Pro Ala Ile
140             275            280            285
141     Gly Lys Thr Glu Lys Thr Gly Lys Arg Ser Thr Ser Ala Ala Val Leu
142             290            295            300
143     Glu Ala Leu Arg Glu Ala His Pro Ile Val Glu Lys Ile Leu Gln Tyr
144             305            310            315            320
145     Arg Glu Leu Thr Lys Leu Lys Ser Thr Tyr Ile Asp Pro Leu Pro Asp

```

## RAW SEQUENCE LISTING

DATE: 05/09/2003

PATENT APPLICATION: US/09/870,353A

TIME: 14:31:42

Input Set : N:\Crf4\05062003\I870353A.raw

Output Set: N:\CRF4\05072003\I870353A.raw

```

146          325          330          335
147      Leu Ile His Pro Arg Thr Gly Arg Leu His Thr Arg Phe Asn Gln Thr
148          340          345          350
149      Ala Thr Ala Thr Gly Arg Leu Ser Ser Asp Pro Asn Leu Gln Asn
150          355          360          365
151      Ile Pro Val Arg Thr Pro Leu Gly Gln Arg Ile Arg Arg Ala Phe Ile
152          370          375          380
153      Ala Glu Glu Gly Trp Leu Leu Val Ala Leu Asp Tyr Ser Gln Ile Glu
154      385          390          395          400
155      Leu Arg Val Leu Ala His Leu Ser Gly Asp Glu Asn Leu Ile Arg Val
156          405          410          415
157      Phe Gln Glu Gly Arg Asp Ile His Thr Glu Thr Ala Ser Trp Met Phe
158          420          425          430
159      Gly Val Pro Arg Glu Ala Val Asp Pro Leu Met Arg Arg Ala Ala Lys
160          435          440          445
161      Thr Ile Asn Phe Gly Val Leu Tyr Gly Met Ser Ala His Arg Leu Ser
162          450          455          460
163      Gln Glu Leu Ala Ile Pro Tyr Glu Glu Ala Gln Ala Phe Ile Glu Arg
164      465          470          475          480
165      Tyr Phe Gln Ser Phe Pro Lys Val Arg Ala Trp Ile Glu Lys Thr Leu
166          485          490          495
167      Glu Glu Gly Arg Arg Arg Gly Tyr Val Glu Thr Leu Phe Gly Arg Arg
168          500          505          510
169      Arg Tyr Val Pro Asp Leu Glu Ala Arg Val Lys Ser Val Arg Glu Ala
170          515          520          525
171      Ala Glu Arg Met Ala Phe Asn Met Pro Val Gln Gly Thr Ala Ala Asp
172      530          535          540
173 Leu Met Lys Leu Ala Met Val Lys Leu Phe Pro Arg Leu Glu Glu Met
174      545          550          555          560
175      Gly Ala Arg Met Leu Leu Gln Val His Asp Glu Leu Val Leu Glu Ala
176          565          570          575
177      Pro Lys Glu Arg Ala Glu Ala Val Ala Arg Leu Ala Lys Glu Val Met
178          580          585          590
179      Glu Gly Val Tyr Pro Leu Ala Val Pro Leu Glu Val Glu Val Gly Ile
180          595          600          605
181      Gly Glu Asp Trp Leu Ser Ala Lys Glu Gly Ile Asp Gly Arg Gly Gly
182          610          615          620
183      Gly Gly His His His His His His
184      625          630
186 <210> SEQ ID NO: 5
187 <211> LENGTH: 2763
188 <212> TYPE: DNA
189 <213> ORGANISM: Artificial Sequence
190 <220> FEATURE:
191 <223> OTHER INFORMATION: Description of Artificial Sequence:fusion protein
192      Ssod7d/full-length Tag
193 <220> FEATURE:
194 <221> NAME/KEY: CDS
195 <222> LOCATION: (1)..(2763)

```

## RAW SEQUENCE LISTING

DATE: 05/09/2003

PATENT APPLICATION: US/09/870,353A

TIME: 14:31:42

Input Set : N:\CrF4\05062003\I870353A.raw

Output Set: N:\CRF4\05072003\I870353A.raw

```

196 <223> OTHER INFORMATION: Ssod7d/full-length Tag
197 <400> SEQUENCE: 5
198   atgattacga attcgagcgc aaccgtaaaag ttcaagtaca aaggcgaaga aaaagaggta 60
199   gacatctcca agatcaagaa agtatggcgt gtgggcaaga tgatctcctt cacctacgac 120
200   gagggcggtg gcaagaccgg ccgtgggtgcg gtaagcgaaa aggacgcgcc gaaggagctg 180
201   ctgcagatgc tggagaagca gaaaaagggc ggcgtgtca ctagtggat gctgcccctc 240
202   tttgagccca agggccgggt cctcctgggt gacggccacc acctggccta ccgcaccttc 300
203   cacgccctga agggcctcac caccagccgg ggggagccgg tgcaggcgggt ctacggcttc 360
204   gccaaagacc tcctcaaggc cctcaaggag gacggggacg cggatgatcgt ggtctttgac 420
205   gccaaaggcc cctccttcgg ccacgaggcc tacgggggggt acaaggcggg ccggggcccc 480
206   acgccagagg actttccccg gcaactcgcc ctcatcaagg agctggtgga cctcctgggg 540
207   ctggcgcgcc tcgaggtccc gggctacgag gcggacgacg tcctggccag cctggccaag 600
208   aaggcggaag aggagggcta cgaggtccgc atcctcaccg ccgacaaaag cctttaccag 660
209   ctctttccg accgcatcca cgtcctccac cccgaggggt acctcatcac cccggcctgg 720
210   ctttgggaaa agtacggcct gagggccgac cagtgggccc actaccgggc cctgaccggg 780
211   gacgagtcgg acaaccttcc cggggtcaag ggcatcgggg agaagacggc gaggaagctt 840
212   ctggaggagt gggggagcct ggaagccctc ctcaagaacc tggaccggct gaagcccgcc 900
213   atccgggaga agatcctggc ccacatggac gatctgaagc tctcctggga cctggccaag 960
214   gtgcgcaccg acctgcccct ggaggtggac ttcgccaaaa ggcgggagcc cgaccgggag 1020
215   aggcctaggg cctttctgga gaggccttag tttggcagcc tcctccacga gttcggcctt 1080
216   ctggaagacc ccaaggccct ggagggagcc ccttgcccc cgccggaagg ggccttcgtg 1140
217   ggctttgtgc tttccgcaa ggagcccatc tgggcccagc ttctggccct ggccgcccgc 1200
218   agggggggcc ggggccaccg ggcccccgag ccttataaag ccctcaggga cctgaaggag 1260
219   gcgcgggggc ttctcgccaa agacctgagc gttctggccc tgagggaagg ccttggcctc 1320
220   ccgcccggcg acgaccccat gctcctcgcc tacctcctgg accttccaa caccaccccc 1380
221   gagggggtgg cccggcgcta cggcggggag tggacggagg aggcggggga gcggggccgc 1440
222   ctttccgaga ggctcttcgc caacctgtgg gggaggcttg agggggagga gaggtcctt 1500
223   tggctttacc gggaggtgga gagggccctt tccgctgtec tgggccacat ggaggccacg 1560
224   ggggtgcgcc tggacgtggc ctatctcagg gccttgctcc tggaggtggc cgaggagatc 1620
225   gccgcctcg aggcgaggt cttccgctg gccggccacc ccttcaacct caactcccgg 1680
226   gaccagctgg aaagggtcct ctttgacgag ctagggttc ccgccatcg caagacggag 1740
227   aagaccggca agcgtccac cagcgccgcc gtcctggagg ccctccgga ggcccacccc 1800
228   atcgtggaga agatcctgca gtaccgggag ctaccaagc tgaagagcac ctacattgac 1860
229   cccttgccgg acctcatcca ccccaggagc ggccgcctcc acaccgctt caaccagacg 1920
230   gccacggcca cgggcaggct aagtagctcc gatcccaacc tcagaaacat ccccgctccg 1980
231   accccgcttg ggagaggat ccgcccggcc ttcacgcggc aggaggggtg gctattggtg 2040
232   gccctggact atagccagat agagctcagg gtgctggccc acctctccgg cgacgagaac 2100
233   ctgatccggg tcttcagga gggcgggac atccacagc agaccgcag ctggatgttc 2160
234   ggcttcccc gggagggcgt ggacccccgt atgcgcggg cgccaagac catcaacttc 2220
235   ggggtcctct acggcatgtc ggcccaccgc ctctcccagg agctagccat cccttacgag 2280
236   gagggccagg ccttcattga gcgctacttt cagagcttcc ccaaggtgcy ggcttgatt 2340
237   gagaagacct tggaggagg caggaggcgg gggtagctgg agacctctt cgccgccgc 2400
238   cgctacgtgc cagacctaga ggcccgggtg aagagcgtgc gggaggcggc cgagcgcatg 2460
239   gccttcaaca tgcccgcca gggcaccgcc gccgacctca tgaagctggc tatggtgaag 2520
240   ctcttcccc ggctggagga aatgggggccc aggatgctcc ttcaggtcca cgacgagctg 2580
241   gtcctcgagg ccccaaaaag gagggcgag gccgtggccc ggctggccaa ggaggtcatg 2640
242   gagggggtgt atcccctggc cgtgcccctg gaggtggagg tggggatagg ggaggactgg 2700
243   ctctccgcca aggagggcat tgatggccgc ggcggaggcg ggcatcatca tcatcatcat 2760
244   taa 2763

```

**VERIFICATION SUMMARY**

DATE: 05/09/2003

PATENT APPLICATION: US/09/870,353A

TIME: 14:31:43

Input Set : N:\Crf4\05062003\I870353A.raw

Output Set: N:\CRF4\05072003\I870353A.raw

L:7 M:270 C: Current Application Number differs, Wrong Format